Machine Learning Based Mathematical Modeling of Heat Treatment on Nanocoated Glass Systems

BENAN AKCA, National Defense University, Naval Academy, Turkey
SENIZ TURKUZ, Science and Technology Center, Şişecam Inc., Turkey
SINEM ERASLAN, Science and Technology Center, Şişecam Inc., Turkey
BATUHAN GUNDOGDU∗, National Defense University, Naval Academy, Turkey

The low-emission (Low-E) glass products that meet the needs of the construction and automotive industry are usually obtained by depositing virtually invisible nano-scale multilayer structures on the glass. Within the scope of safety regulations, tempering is carried out after deposition. On the other hand, the optical spectra curves of these coated glasses which give information about the transmission and reflectance properties of the product, exhibit change as a result of the tempering process. Thus, determining the change after the tempering process becomes an important point during the product development stages. Since there is no analytical method for calculating the effect of heat treatment, it can only be roughly predicted by the information obtained from the previous heat treatment applications which require considerable time and cost. Therefore, it becomes imperative to model the effect of heat treatment on optical spectra for practical manufacturing of desired optical quality. In this paper, heat treatment effects, which change the optical spectra characteristics of low emissivity glass is modeled by machine learning techniques, by training neural networks feeding the pre and post-heating measurements through their input and output layers. Experiments conducted on Low-E glass coatings exhibit the efficiency of this method on several datasets, with various coating architectures.

CCS Concepts: • Computing methodologies → Supervised learning by regression; Neural networks; Model development and analysis; Model verification and validation.

Additional Key Words and Phrases: function learning, heat treatment, neural networks, nanocoated glass systems, optical spectra, process modeling

ACM Reference Format:

1 INTRODUCTION

In recent years, coated glass has emerged as a product group that has taken part in the technology race in the automotive, architectural and display industries. With the development of technology and increasing industrial requirements, increasingly complex coatings have been produced. Some

∗Corresponding author.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

© 2019 Association for Computing Machinery.
0004-5411/2019/10-ART111 $15.00
https://doi.org/10.1145/1122445.1122456
of those; Indium tin oxide (ITO) coatings for LCDs, Low-E coatings used to minimize heat loss indoors and solar control coatings used to reflect some of the sun’s heat to maintain the heat degree of indoor [3]. Within the scope of safety regulations, the coated glass used in industrial areas where safety is at the forefront, such as the automotive industry, is tempered after coating. Tempered glass is a glass that is critically heated and suddenly cooled to increase strength. These tempered glasses are more secure than untreated glasses because they become more durable and will be broken down into small pieces in case of a breakage [7]. Thanks to these properties, the tempering process is also applied to coated glass. In the output of the coated glass which has completed the heat treatment process, it is necessary to determine how the optical performance properties of the glass will be affected during the heat treatment to have the desired optical performance. Various analytical methods have been used to predict the optical properties of thin films that have not yet been tempered [4, 9, 12, 20]. However, the optical spectrum curves, which contain important characteristics such as reflectance/transmittance performance and color of Low-E coated glasses, show changes that cannot be modeled as a result of various chemical and physical interactions on the film after heat treatment. This has led to the need to model the effect of heat treatment on optical spectra in coated glass systems by another method.

Modeling of processes that cannot be modeled analytically by machine learning (ML) is one of the recent studies in the literature. If a sufficient number of input and output examples are obtained for the problem encountered, a mathematical function can be fitted using machine learning models (artificial neural networks, linear models, etc.). When the learning process is done using the past data, outputs for new data can be predicted using this model [1]. Several applications of ML to material science for mathematical modeling have been studied over the past decade. With machine learning techniques; explosion effect modeling [10], friction characteristic modeling [5], defective product material learning relationship [17], modeling of molecular electrical properties [15], optimization of biomedical glass coatings [21], wear modeling [2] can be defined as examples such studies in material science with machine learning techniques.

Specifically, there are a number of studies on the modeling of heat treatment by machine learning techniques within the scope of materials science. Song and Zhang developed an artificial neural network (ANN) to model the relationship between pre-heating deformation, solid solution and aging time property variables of high strength aluminum alloys, final tensile strength and yield strength outputs after heat treatment [19]. They concluded that ANN’s are a viable method for modeling the effects of heat treatment on the mechanical properties of aluminum alloy. Malinov et al. employed an ANN model for the analysis and prediction of the correlation between heat treatment parameters and mechanical properties in titanium alloys [13]. As a result of the training, high success has been achieved in the prediction of properties of titanium alloys at different temperatures as functions of processing parameters and heat treatment cycle. In another material science study using ANN for modeling the heat treatment process, Zanuncia et al. predicted the physical, mechanical and colorimetric properties of the heat-treated eucalyptus grandis tree over 95% success [22]. These studies provide important preliminary information about the feasibility of mathematical modeling of heat treatment in material science.

In this study, the change in the optical properties of the coated glass as a response to the heat treatment applied for tempering is modeled by machine learning techniques. As in the above-mentioned material science studies, the pre and post-heat treatment values of the materials’ (Low-E glass) properties were used to train proposed machine learning models. Changes in optical spectra after heat treatment were predicted with high generalization and accuracy by trained models in various experiments.

The contributions of this study can be expressed as follows;
Machine learning-based mathematical modeling of heat treatment of nanocoated glass systems has been introduced. To our knowledge, this is the first study to attempt and successfully fulfill this gap in the industry.

The effects of using coating properties in training in addition to optical spectrum and wavelength values are investigated.

The effects of using additional synthetic data with high resolution in a specific spectrum range in the training of the model are examined.

The remainder of the paper is organized as follows. In Section 2, we propose a method to enable the use of optical data of coated glass in model training. In Section 3 several experiments and analyzes results are exhibited in the following order;

1. The feasibility of mathematical modeling of the effects of heat treatment is investigated with the basic models: multivariate linear regression (MLR) and ANN using the optical spectrum and wavelength values as input.
2. In addition to optical spectra and wavelength values, the effects of using glass coating properties in model training were examined.
3. The effect of adding high-resolution synthetic data on local performance was analyzed.

In Section 4 an overall conclusion is presented.

As a result of the study, firstly it is concluded that the heat treatment process in the production of coated glasses can be successfully modeled with ANN. Then it is observed that the use of glass coating properties in the model training as an additional feature increased the performance of the model so that the data number can be increased by using the data in different coating properties together. Finally, the addition of high-resolution local synthetic data drastically improved model performance for the corresponding local spectra.

2 MATERIAL AND METHODOLOGY

2.1 Material

Three sets of datasets were used to experiment with the methodology proposed in this paper. Since we aim to model the optical effects of heat treatment, we collect optical measurements for various glass coatings before and after the heating process. The optical measurements we use in this study are transmittance ($T$), the reflectance of coated surface ($R_c$) and reflectance of uncoated surface ($R_u$).

In the datasets created within the scope of this study, glasses used in the experiments were coated with Low-E technology in Šişecam Science and Technology Center R & D laboratories. In the experiments carried out, a specific type of Low-E coating was studied on all glasses (Fig. 1). Optic measurements were obtained by spectrophotometer from each sample in the $280nm$ – $2500nm$ range and $5nm$ resolution. Each sample has separate optical spectra for its $T$, $R_c$ and $R_u$ optical properties.
Physical and chemical changes occur in the structure of the material as a result of exposure to very high temperatures and sudden cooling during the tempering process. As a result of this
structural change in glass and coating film, the characteristics of the reflection and transmittance optical spectra of the product exhibit drastic changes. It can be observed in Fig. 2 that this change occurs most intensely in the visible spectrum (380 – 780 nm).

2.2 Methodology

In this study, supervised machine learning techniques have been developed for the prediction of the optical spectrum formed after heat treatment. The purpose of supervised learning is to obtain a function to explain a relationship between the input and the output using input-output pairs that have been experienced in the context of the problem [18]. As in the description, we used the pre and post-heat treatment optical values obtained from the previous processes to obtain a function explaining the effect of the heat treatment process.

2.2.1 Model development for machine learning. As the light passes through the glass, the optical properties of the glass vary depending on the wavelength of the light exposure. One of the methods that could be used for modeling the input-output relationship of heat treatment is to input the optical property values corresponding to each wavelength along with the optical spectra as input to the machine learning model as shown in Fig. 3. In this case, since the number of parameters at the input of the machine learning model will be as high as the number of optical measurements, the number of sample pairs used on training must be sufficient to prevent overfitting. However, since the optical spectra of each sample can be evaluated as a single sample pair with this method, the amount of data available would be inevitably inadequate for this method to yield efficient results (high generalization and low error).

Therefore, instead of using the entire spectra as input parameters, optical property values based on wavelengths were used along with the wavelength values itself, as input parameters. This arrangement could be considered as a practical data augmentation methodology and is one of the key contributions of this paper. In other words, the function \( f(\cdot) \) is expected to predict the optical values after the heat treatment based on the input optical values and wavelength values. Mathematically this desired function \( f(\cdot) \) has 4 inputs and 3 outputs and can be shown as follows:

\[
\begin{bmatrix}
\hat{T}_{\text{post}}(\lambda), \\
\hat{R}_{\text{post}}^c(\lambda), \\
\hat{R}_{\text{post}}^u(\lambda)
\end{bmatrix} = f(\lambda, T_{\text{pre}}(\lambda), R_{\text{pre}}^c(\lambda), R_{\text{pre}}^u(\lambda))
\]  

Where:
- \( \lambda \) is the wavelength of light that varies between 280nm – 2500nm.
- \( T_{\text{pre}} \) is the transmittance measurement before heat treatment.
- \( R_{\text{pre}}^c \) is the coated side reflection measurement before heat treatment.
- \( R_{\text{pre}}^u \) is the uncoated side reflection measurement before heat treatment.

Fig. 3. A viable yet impractical method for modeling the heat treatment process.
• \( \hat{T}_{\text{post}} \) is the prediction of transmittance after heat treatment.
• \( \hat{R}_{\text{c post}} \) is the prediction of coated side reflection after heat treatment.
• \( \hat{R}_{\text{u post}} \) is the prediction of uncoated side reflection after heat treatment.

The hat symbols in the representations \( \hat{T}_{\text{post}} \), \( \hat{R}_{\text{c post}} \) and \( \hat{R}_{\text{u post}} \) indicate that these values are predicted.

With the augmentation brought by this methodology, the number of input parameters can be reduced significantly. We would like to take the liberty to give specific numbers from our experiments in this methodology section, in order to emphasize the mathematical effect of this data in terms of the curse of dimensionality [1]. Our measurements were taken between 280 – 2500 nm with 5 nm resolution, so each spectrum has 445 discrete \( T \), \( R_c \) and \( R_u \) values. Therefore the number of input parameters reduced by 333 times (from 3x445 to 4) compared to the first method, while the number of sample pairs can be increased by 445 times.

The input-output relationship described above is given in Fig. 4 on the proposed machine learning model.

![Fig. 4. Recommended mathematical method for learning the effect of heat treatment.](image)

2.2.2 **Multivariate linear regression.** Regression analysis is a statistical technique used to investigate and model the relationship between parameters [16]. Function models that explain the relationship between variables linearly are called linear regression. Within the scope of the study, we can express the input-output relationship in the proposed model as linear with Eq. 2.

\[
y_1 = a_{i0} + a_{i1}\lambda + a_{i2}T_{\text{pre}}(\lambda) + a_{i3}R_{c \text{pre}}(\lambda) + a_{i4}R_{u \text{pre}}(\lambda)
\]

Where \( y_1 \), \( y_2 \) and \( y_3 \) refer to \( \hat{T}_{\text{post}}(\lambda) \), \( \hat{R}_{c \text{post}}(\lambda) \) and \( \hat{R}_{u \text{post}}(\lambda) \) respectively.

2.2.3 **Artificial neural network.** ANNs are one of the most used methods in machine learning problems are designed with inspiration from the complex neuronal networks of biological learning systems [8]. As in natural neural networks, in ANN, the data from the output of different neurons are processed by the linked neurons and delivered to the processor neuron’s output.

The general ANN model, which is formed within the scope of the problem addresses in this paper is shown in Fig. 5. The input stage is the pre-heat feature vector defined in Eq. (1). The hidden layer is found by a linear matrix operation applied to the input layer. The width of each layer (\( N_l \) value in Fig. 5) or model’s depth (\( L \) value in Fig. 5) can be changed as a design parameter according to the complexity of the process to be modeled.
Based on this definition, the expression of the hidden layer is as follows:

\[ h_l = W_l a_{l-1} + b_l \]  

(3)

Where \( l \) is the layer number, \( h \) is the output of the linear operation, \( W \) is the weight matrix, \( a \) is the output of activation function on \( l - 1 \) and the input vector of \( l \), \( b \) represents the bias vector of \( l \). For the first layer \( (l = 1) \) operations the feature vector \((x)\) is used instead of \( a_0 \).

After this linear operation, each dimension of the hidden layer is subjected to activation function operation to obtain non-linearity. In this study, tangent hyperbolic function was used as the activation function (Eq.(4)).

\[ a_l(i) = \tanh(h_l(i)) = \frac{e^{h_l(i)} - e^{-h_l(i)}}{e^{h_l(i)} + e^{-h_l(i)}} \]  

(4)

The output layer is another linear layer used to obtain the three-dimensional vector containing the predictions for \( T \), \( R_c \) and \( R_u \). For the ANN model, the non-linear function passed through the penultimate hidden layer is multiplied by the last layer’s matrix to reach the output without being activated. If we call the three rows of \( W_L \) as \( v_1, v_2 \) and \( v_3 \) we can denote the predictions as follows:

\[ \hat{T}_{\text{post}} = v_1 a_L + b_1 \]
\[ \hat{R}_{\text{post}} = v_2 a_L + b_2 \]
\[ \hat{R}_{\text{post}} = v_3 a_L + b_3 \]  

(5)

The parameters, i.e the weight matrices of the ANN are learned by the error backpropagation method. The backpropagation method uses the gradient descent algorithm to optimally adjust the network parameters to the sample pairs in the training set [14].
2.2.4 Performance criteria. The accuracy of the predictions is evaluated by the squared error between the prediction and the actual post-heating values. Eq. (6) was used to measure the MSE value of the test set containing N samples.

In the experiments, instead of separating the datasets into training, development and test sets, we separated them only as training and test sets to use our limited data more effectively. In doing so, we have repeated the experiment several times \( r = 1 \cdots R \) in order not to waive the measurement of the generalization feature of the model. For each repetition, we rearranged the training and test sets and made sure that the data used in the training set were not used among the test set samples \( t = 1 \cdots N \).

\[
MSE_r(T) = \frac{1}{N S} \sum_{t} \sum_{\forall \lambda} (t^T_{\text{post}}(\lambda) - \hat{t}^T_{\text{post}}(\lambda))^2
\]

\[
MSE_r(R_c) = \frac{1}{N S} \sum_{t} \sum_{\forall \lambda} (t^R_{\text{c post}}(\lambda) - \hat{t}^R_{\text{c post}}(\lambda))^2
\]

\[
MSE_r(R_u) = \frac{1}{N S} \sum_{t} \sum_{\forall \lambda} (t^R_{\text{u post}}(\lambda) - \hat{t}^R_{\text{u post}}(\lambda))^2
\]

(6)

Where;

- \( T_{\text{post}} \) is the observed transmittance after heat treatment.
- \( R_{\text{c post}} \) is the observed coated side reflection after heat treatment.
- \( R_{\text{u post}} \) is the observed uncoated side reflection after heat treatment.
- \( \hat{T}_{\text{post}} \) is the prediction of transmittance after heat treatment.
- \( \hat{R}_{\text{c post}} \) is the prediction of coated side reflection after heat treatment.
- \( \hat{R}_{\text{u post}} \) is the prediction of uncoated side reflection after heat treatment.
- \( \lambda \) is the wavelength of light that varies between 280nm – 2500nm.
- \( r \) is the repetition number of the experiment : \( r = 1 \cdots R \)
- \( t \) is the sample number in the test set : \( t = 1 \cdots N \)
- \( S \) is the total number of the wavelengths determined by the resolution used in the model.
- \( N \) represents the total number of samples in the test set.

The average of MSE (\( \overline{MSE} \)) of the test sets was calculated using Eq. (7) to determine the overall performance of the model.

\[
\overline{MSE} = \frac{1}{R} \sum_{r=1}^{R} MSE_r
\]

where,

\[
MSE_r = MSE_r(T) + MSE_r(R_c) + MSE_r(R_u)
\]

(8)

An experiment may have considerably good results in several predictions, and these results may have decreased the \( \overline{MSE} \) of the experiment, or the opposite. With this intuition, it is useful to monitor an additional metric to evaluate the model. In this context, the standard deviation calculation, in which the generalization of the model can be evaluated, was made over MSE values of different combinations as below:

\[
\sigma_{MSE} = \sqrt{\frac{1}{R-1} \sum_{r=1}^{R} (MSE_r - \overline{MSE})^2}
\]

(9)
3 EXPERIMENTAL RESULTS
3.1 Data preparation

Three datasets have been created to be used in different experiments performed in the scope of modeling and performance enhancement.

- **Dataset-I** (DS-I) consists of samples of glass coatings whose physical properties do not vary.
- **Dataset-II** (DS-II) was used in experiments to evaluate the generalization power of the solution. For this reason, DS-II includes samples of glass coatings whose physical properties vary among themselves and include much more data than DS-I. In this dataset, the total NiCr thickness is kept the same, for it is observed to have the most significant effect on change in optical spectra due to the heat treatment.
- **Dataset-III** (DS-III) was used to investigate the effects of the coating-dependent features. In this dataset, we also varied the NiCr thicknesses to develop our models further for better generalization. In this context, in addition to the $T$, $R_c$ and $R_u$ optic values included in the other datasets, in DS-III, NiCr coating thickness data are also used. The structure of datasets is shown in Table 1.

Since the feature vectors $(\lambda, T^{pre}(\lambda), R_c^{pre}(\lambda), R_u^{pre}(\lambda))$ are on different scales were normalized to have zero mean and unity variance to facilitate the learning (Eq. 10).

$$\hat{x}_i = \frac{x_i - \mu_{x_i}}{\sigma_{x_i}} \quad (10)$$

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Total Number of Glass Coatings</th>
<th>Training Set</th>
<th>Test Set</th>
<th>Sample Resource</th>
<th>NiCr Layer Thickness</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>6</td>
<td>5</td>
<td>1</td>
<td>Kept same</td>
<td>Kept same</td>
</tr>
<tr>
<td>II</td>
<td>36</td>
<td>30</td>
<td>6</td>
<td>Varied</td>
<td>Kept same</td>
</tr>
<tr>
<td>III</td>
<td>18</td>
<td>15</td>
<td>3</td>
<td>Varied</td>
<td>Varied</td>
</tr>
</tbody>
</table>

3.2 ANN model of heat treatment

As a baby step to test our methodology, a linear regression model was trained on the DS-I, which includes samples from a dataset whose physical properties were kept the same over the set. As seen in Table 1, 5 samples were used in training and the remaining post-heating spectrum was predicted. The predicted values were compared with the real values of tempered coated glass.

For all experiments executed in this study Keras machine learning framework [6] were used for training and prediction processes. As a result of several training and prediction series, it was observed that the linear regression approach is reasonable but lacking sufficient complexity in explaining the heat treatment process. Therefore, the effect of heat treatment on optical spectra was modeled by ANN using the same dataset. After several epochs of training loop with the ANN model, much more successful results were obtained than the linear model, due to its non-linearity feature. The drastic drop in $\overline{MSE}$ when the nonlinear model used, can be seen in Table 2 for DS-I experiments.
Table 2. Experimental Details.

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Model</th>
<th>No. of hidden layers</th>
<th>No. of nodes</th>
<th>Activation function</th>
<th>Optimization algorithm</th>
<th>Mini batch dimension</th>
<th>No. of epochs</th>
<th>No. of combinations (R)</th>
<th>MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>DS-I</td>
<td>Linear</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Gradient Descent</td>
<td>1</td>
<td>40</td>
<td>15</td>
<td>11.3</td>
</tr>
<tr>
<td>DS-I</td>
<td>ANN-I</td>
<td>1</td>
<td>4-40-3</td>
<td>tanh</td>
<td>ADAM[11]</td>
<td>32</td>
<td>40</td>
<td>15</td>
<td>0.99</td>
</tr>
<tr>
<td>DS-II</td>
<td>ANN-II</td>
<td>2</td>
<td>4-70-70-3</td>
<td>tanh</td>
<td>ADAM[11]</td>
<td>32</td>
<td>100</td>
<td>30</td>
<td>1.48</td>
</tr>
</tbody>
</table>

Throughout the experiments, we worked on three optical spectra, these are; the actual post-heat spectrum, the predicted post-heat spectrum and the pre-heat spectrum. Our ultimate goal is to obtain the resemblance of the predicted spectrum as much as possible to the real spectrum after heat treatment by training the model with a pre-heat spectrum. It is seen in the three spectra given in Fig. 6 that the predicted curve of the test set almost exactly matches the actual optical spectra curve. Such promising results led us to think that the heat treatment process of coated glass could be modeled with ANN. We continued the experiments with different architectures and datasets to assess the generalization power of the proposed model and feasibility on different glass coatings.

Fig. 6. Comparison of predicted optical spectra using linear regression (left) and ANN (right).

3.2.1 Testing the generalization of the model. As seen in the previous experiment, the ANN approach to the heat treatment modeling problem yields successful results. However, experiments with samples obtained from DS-I whose physical properties do not vary along the samples, fail to provide a proof for the generalization of the model. Therefore, it can not provide an adequate measurement to determine the final performance. For this reason, the model was further evaluated
using DS-II that has 36 samples in which the physical properties and layer characteristics vary between samples. In this context, a deeper architecture ANN, denoted ANN-II and whose structure is shown in Table 2, was used. It is now a good point to further explain the experimental details given with Table 1 and 2. We randomly segmented the Datasets into two parts as given in Table 1 and we conducted this random segmentation $R$ times, to account for and eliminate the effects of chance in evaluating our results. Therefore for ANN-II, we tested on 6 glass coatings and did this randomly on 30 different test-set combinations which can be seen from Table 1 and 2, respectively.

To facilitate the observation of the generalization success in different glasses more easily, 3 samples with the most aggressive optical characteristics were selected from DS-II. We intentionally chose these samples that exhibit the most drastic change as a response to heat treatment, to exhibit the mathematical model behavior. Post-heat transmittance values ($T$) of these samples were predicted with different models. Both inter-model prediction performance and the performance of each model in different glass characters are shown in Fig. 7.

![Fig. 7. Inter-model performance exhibition for three sample.](image)

A deeper examination of the generalization ability of the model posed a question “Could this model be successful in a dataset that includes a different coating property, such as the coating thickness of NiCr layer?” If the model were to remain successful despite the varying thickness values of NiCr, it would indicate that our model had a high level of generalization and that the dataset could be increased with samples containing different layer thicknesses. Therefore, we continued the experiments with DS-III which involves different coating thickness data to examine the answer to this question.

3.2.2 Sample specific features: NiCr layer’s thickness as a feature. On the development and initial experiments, we held a very simply put, yet considerably brave assumption: “The heat treatment effect is a non-linear function effective on optical characteristics of a nanocoated glass, only dependent
on such characteristics of the coating.” The simplicity of the assumption did in fact brought the desired modeling performance, yet now in this section, we aim to improve the model by incorporating a coating-depending feature as input to the function to be learned.

In the datasets, each sample contains optical property values corresponding to 445 wavelengths. Different samples exhibit varying optical behavior for each wavelength. The more variable the optical value of the samples in the dataset corresponding to the same wavelength, the more difficult the prediction is made with the trained model. To measure this variability, we calculated the standard deviation of the transmittance values of the samples corresponding to the same wavelength. Fig. 8 shows the standard deviation estimations for each wavelength calculated over the samples on each dataset. It can be observed that DS-III is the most aggressive dataset where the samples show the most different characteristics, as observed by high variance.

![Fig. 8. Comparison of standard deviation values of transmittance feature before heat treatment for each wavelength value.](image)

Predicting post-heat optical values on such an aggressive dataset with high accuracy indicates a relatively high generalization feature compared to previous models. To contribute to the training of the model in such conditions we added NiCr layer thickness as another feature to the proposed model’s input layer.

To avoid confusion, we have continued to call the models used in this experiment with their names before the change, since only the number of input features changed and the other architectural features are kept same. As a result of using NiCr thickness as a feature, a high level of success was achieved with low error on predicting in the data which had never been used in training before. The use of samples of different NiCr thicknesses together increased the diversity of the dataset, thus increasing the amount of the sample pair, which is one of the factors directly affecting the prediction success of the ANN model. This case resulted in the idea of combining different datasets (combined dataset).

Based on this idea, we also investigated using DS-II, whose NiCr thickness was kept same, in addition to DS-III to observe the effects of increasing the data size and monitor the prospective improvements as the actual measurements take place in the future. Table 3 summarizes the effects of using NiCr as feature and the data combination.

J. ACM, Vol. 37, No. 4, Article 111. Publication date: October 2019.
Table 3. Performance results according to models and datasets (MSE).

<table>
<thead>
<tr>
<th></th>
<th>DS-II</th>
<th>DS-III</th>
<th>DS-III (NiCr used as parameter)</th>
<th>Combined Dataset</th>
<th>Combined Dataset (NiCr used as parameter)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN-I</td>
<td>1.482</td>
<td>2.098</td>
<td>2.258</td>
<td>1.267</td>
<td>1.253</td>
</tr>
<tr>
<td>ANN-II</td>
<td>1.495</td>
<td>1.507</td>
<td>1.505</td>
<td>1.266</td>
<td>1.209</td>
</tr>
</tbody>
</table>

3.2.3 Local performance improvement with synthetic data. After boosting the performance of our model we aimed for the local performance improvement. When we look at Fig. 2 what is striking is, at the visible spectra (380 – 780 nm) optic values oscillation is more than the rest of all. Therefore the visible light region is the most difficult region for prediction because of high-level oscillation. In order to focus the model on this relatively difficult region to predict, we used additional synthetic data in the training of the model.

In the datasets used in this study, 445 optical values were obtained with 5 nm sampling period from spectrum containing 280 nm – 2500 nm wavelength range of each glass sample. This process was done separately for each of the three optical features (T, Rc and Ru). By using these real values obtained by spectrophotometer, intermediate values were synthetically estimated at 1 nm resolution by interpolation for the visible light region. Thus, in addition to the actual values between 380 – 780 nm, nearly true synthetic values were obtained. More specifically we used 3x402 synthetic data (sample x spectral range) containing 1 nm resolution wavelength in addition to 30x445 data containing 5 nm resolution wavelength for the target wavelength range (380 – 780 nm) in training for improving the performance of the ANN model in the targeted region. As a result of the experiment, localized performance improvement on the visible light region was obtained as seen in Fig. 9.

![Fig. 9. Comparison of the predicted transmittance spectrum with the general model (left) and the local performance improvement model (right) which use additional high-resolution synthetic data in training.](image)

4 CONCLUSION

In this paper, we obtained a mathematical function to explain the changes that occur on optical spectra of nanocoated glass systems as a response to heat treatment. We followed a machine learning approach and evaluated the performance of different models.
learning based approach using the past measurements of the optical spectra before and after the heat treatment in the artificial neural network training, which in turn, yielded a model to successfully estimate the optical spectra of heating-applied glass coatings that have not been used in model training. To our knowledge, this study the first attempt to address this issue with this method. We have shown that such a technique yields a feasible and practical means to estimate the effects of heat treatment. We observed that nonlinearity is required to model the chaotic phenomenon of heat treatment on optical spectra and that the generalization ability of the model decreases with the complexity of the model in datasets of varying optical characteristics. The desired non-linearity is obtained by multi-layer artificial neural networks. We have examined three solutions to address the conundrum of generalization:

1. Increasing the number of sample pairs used in training by combining datasets,
2. Using coating properties in the training to reinforce the model.
3. Improving performance in difficult-to-predict areas by using additional synthetic data.

As future study, different layer thickness values can also be incorporated into the model, such as Ox and dielectric layers. Order of the coating layers can be used as parameters or a generative model can be used to learn proper coating structure for desired optical behavior.

REFERENCES

